

EAST Search History

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	1467	(715/500).CCLS.	US-PGPUB; USPAT; JPO; DERWENT	OR	OFF	2007/11/12 10:30
L2	1188	L1 and (@ad<"20031015" or @rlad<"20031015")	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/11/12 10:31
L3	0	L2 and (power\$point and OCR and (markup NEAR2 language) and raster)	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/11/12 10:31
L4	34	(power\$point and OCR and (markup NEAR2 language) and raster)	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/11/12 10:31
L5	7	L4 and (@ad<"20031015" or @rlad<"20031015")	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/11/12 10:33
L6	2	(OCR SAME (markup NEAR2 language) SAME raster)	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/11/12 10:33
L7	2	L6 and (@ad<"20031015" or @rlad<"20031015")	US-PGPUB; USPAT; JPO; DERWENT	OR	ON	2007/11/12 10:33



USPTO

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

 Search: ☒ The ACM Digital Library ☐ The Guide

data extraction PDF

SEARCH

THE ACM DIGITAL LIBRARY

[Feedback](#) [Report a problem](#) [Satisfaction survey](#)
Terms used: [data extraction](#) [PDF](#)

Found 46,636 of 214,158

Sort results by

relevance

☒ [Save results to a Binder](#)[Try an Advanced Search](#)

Display results

expanded form

☒ [Search Tips](#)[Try this search in The ACM Guide](#)☐ Open results in a new window

Results 1 - 20 of 200

Result page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale ☐ ☐ ☐ ☐ ☐**1** [Variable data printing: Extracting reusable document components for variable data](#)[printing](#)

Steven R. Bagley, David F. Brailsford, James A. Ollis

August 2007 **Proceedings of the 2007 ACM symposium on Document engineering****DocEng '07****Publisher:** ACM PressFull text available: [pdf\(748.64 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Variable Data Printing (VDP) has brought new flexibility and dynamism to the printed page. Every printed instance of a specific class of document can now have different degrees of customized content within the document template.

This flexibility comes at a cost. If every printed page is potentially different from all others it must be rasterized separately, which is a time-consuming process. Technologies such as PPML (Personalized Print Markup Language) attempt to address this problem ...

Keywords: PDF, SVG, content extraction, graphic objects, posrScript, variable data printing

2 [Invited talk: The Lixto data extraction project: back and forth between theory and](#)[practice](#)

Georg Gottlob, Christoph Koch, Robert Baumgartner, Marcus Herzog, Sergio Flesca

June 2004 **Proceedings of the twenty-third ACM SIGMOD-SIGACT-SIGART****symposium on Principles of database systems PODS '04****Publisher:** ACM PressFull text available: [pdf\(430.70 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

We present the Lixto project, which is both a research project in database theory and a commercial enterprise that develops Web data extraction (wrapping) and Web service definition software. We discuss the project's main motivations and ideas, in particular the use of a logic-based framework for wrapping. Then we present theoretical results on monadic datalog over trees and on Elog, its close relative which is used as the internal wrapper language in the Lixto system. These results include both ...

3 [Information extraction 1: Integrating data and text mining processes for digital library](#)[applications](#)

Robert Sanderson, Paul Watry